

Natural scene reconstruction from fMRI signals using generative latent diffusion

Year: 2023 | Citations: 136 | Authors: Furkan Ozcelik, Rufin VanRullen

Abstract

In neural decoding research, one of the most intriguing topics is the reconstruction of perceived natural images based on fMRI signals. Previous studies have succeeded in re-creating different aspects of the visuals, such as low-level properties (shape, texture, layout) or high-level features (category of objects, descriptive semantics of scenes) but have typically failed to reconstruct these properties together for complex scene images. Generative AI has recently made a leap forward with latent diffusion models capable of generating high-complexity images. Here, we investigate how to take advantage of this innovative technology for brain decoding. We present a two-stage scene reconstruction framework called “Brain-Diffuser”. In the first stage, starting from fMRI signals, we reconstruct images that capture low-level properties and overall layout using a VDVAE (Very Deep Variational Autoencoder) model. In the second stage, we use the image-to-image framework of a latent diffusion model (Versatile Diffusion) conditioned on predicted multimodal (text and visual) features, to generate final reconstructed images. On the publicly available Natural Scenes Dataset benchmark, our method outperforms previous models both qualitatively and quantitatively. When applied to synthetic fMRI patterns generated from individual ROI (region-of-interest) masks, our trained model creates compelling “ROI-optimal” scenes consistent with neuroscientific knowledge. Thus, the proposed methodology can have an impact on both applied (e.g. brain–computer interface) and fundamental neuroscience.